15th International Command & Control Research & Technology Symposium

Command & Control in Virtual Environments: Laboratory Experimentation to Compare Virtual with Physical

Richard D. Bergin, Alicemary Aspell Adams, Ramez Andraus Junior, Bryan J. Hudgens, June G. Chinn Yi Lee, Dr. Mark E. Nissen
US Naval Postgraduate School

Topics:
Topic 5: Experimentation and Analysis
Topic 7: C2 Approaches and Organization

* Point of contact
Richard Bergin
US Naval Postgraduate School
Graduate School of Operational & Information Sciences
589 Dyer Road, Room 227A Monterey, CA 93943
Tel: (831) 656-3112, Fax: (831) 656-3679, E-mail: rdbergin@nps.edu

ABSTRACT

Research in command and control is advancing rapidly through a campaign of laboratory experimentation using the ELICIT (Experimental Laboratory for Investigating Collaboration, Information-sharing, and Trust) multiplayer online counterterrorism intelligence game. In most ELICIT experiments, participants play the game through a Web interface and interact with one another solely through textual information exchange. This mirrors in large part the network-centric environment associated with most counterterrorism intelligence work in practice, and it reflects what appears to be a widespread assumption about how to organize such work: in a physically distributed, virtual manner. Such reflection is consistent with considerable research (e.g., in Educational Psychology) prescribing such distributed, virtual environments for work performance. Alternatively, substantial research (e.g., Media Richness Theory) suggests instead that a more personal, physical environment offers potential to improve performance. Hence we have a theoretical conflict with potential to affect how the important work of counterterrorism intelligence is organized. The research described in this article addresses this theoretical conflict through a series of experiments to assess the comparative performance of people working in physical, face-to-face versus textual, virtual environments. Exercising great care to match experiment conditions and control for factors other than task environment, results elucidate important comparative performance effects and suggest compelling follow-on experiments as well as practical implications.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
JUN 2010		**00-00-2010 to 00-00-2010**

4. TITLE AND SUBTITLE
Command & Control in Virtual Environments: Laboratory Experimentation to Compare Virtual with Physical

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School,Graduate School of Operational & Information Sciences,589 Dyer Road,Monterey,CA,93943

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Proceedings of the 15th International Command & Control Research & Technology Symposium, Santa Monica, CA (June 2010).

14. ABSTRACT
Research in command and control is advancing rapidly through a campaign of laboratory experimentation using the ELICIT (Experimental Laboratory for Investigating Collaboration Information-sharing, and Trust) multiplayer online counterterrorism intelligence game. In most ELICIT experiments, participants play the game through a Web interface and interact with one another solely through textual information exchange. This mirrors in large part the networkcentric environment associated with most counterterrorism intelligence work in practice, and it reflects what appears to be a widespread assumption about how to organize such work: in a physically distributed, virtual manner. Such reflection is consistent with considerable research (e.g., in Educational Psychology) prescribing such distributed, virtual environments for work performance. Alternatively, substantial research (e.g., Media Richness Theory) suggests instead that a more personal, physical environment offers potential to improve performance. Hence we have a theoretical conflict with potential to affect how the important work of counterterrorism intelligence is organized. The research described in this article addresses this theoretical conflict through a series of experiments to assess the comparative performance of people working in physical, face-to-face versus textual, virtual environments. Exercising great care to match experiment conditions and control for factors other than task environment, results elucidate important comparative performance effects and suggest compelling follow-on experiments as well as practical implications.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**19**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Keywords: collaboration, ELICIT, experimentation, virtual environment

Acknowledgements: The research described in this article was funded in part by the Command and Control Research Program through the Center for Edge Power at the Naval Postgraduate School.

INTRODUCTION

Modern military organizations have adapted and evolved over many centuries and millennia, respectively. Hierarchical command and control (C2) organizations in particular have been refined longitudinally (e.g., through iterative combat, training and doctrinal development) to become very reliable and effective at the missions they were designed to accomplish. However, recent research suggests that the Hierarchy may not represent the best organizational approach to C2 in all circumstances (Roger B. Mason, University of Surrey, UK, 2007), particularly where the environment is unfamiliar or dynamic. Indeed, alternate, more flexible C2 organizational approaches such as the Edge have been proposed (Lave & Wenger, 1991) to overcome Hierarchy limitations, but the same recent research suggests that the Edge may not represent the best organizational approach to C2 in all circumstances either, particularly where the environment is familiar and stable.

Of course, the Hierarchy and Edge both represent organizational archetypes (Rivkin, 2000), each of which offers considerable latitude in terms of detailed organizational design and customization. For instance, recent research demonstrates further how the performance of both Hierarchy and Edge organizations is sensitive to factors such as network infrastructure, professional competency and other factors that can be affected through leadership, management and investment (Ramo, 2009). With incessant advances in information technology (IT) that appear to be continuing, one may be able to overcome the limitations inherent in Hierarchy, Edge or other organizations or even enable such organizations to adapt—through IT—to shifting conditions.

This notion is fundamental to Network Centric Operations (NCO), where people and organizations operate principally in network-enabled virtual environments as opposed to their physical counterparts. Unfortunately, empirical evidence to support the asserted superiority of NCO remains sparse, and the capability enhancing properties of virtual environments remain more in the domain of lore than empirical assessment. Indeed, drawing from substantial research in both Educational Psychology and media richness, the counter argument that performance in virtual environments will be *worse* than in physical counterparts offers substantial merit and empirical support. Hence we find some controversy between the tenets of NCO and empirical evidence in related fields.

Building upon these separate streams of research, we continue a campaign of experimentation to assess the relative performance of different C2 organizational approaches across a diversity of environments and conditions. In this present study, we investigate explicitly the impact of virtual versus physical environment on organizational performance, focusing first here on the near-ubiquitous Hierarchy, which remains the predominant approach to C2 organization. Specifically, capitalizing upon the excellent internal validity and control available,

we begin a series of laboratory experiments to assess the relative performance of a Hierarchy in virtual versus physical environments.

In the balance of the paper, we draw from the Educational Psychology and media richness literatures to motivate a set of research hypotheses to address the virtual-physical question well. We then detail our research design and report in turn the key findings and results. The paper closes with a set of conclusions, recommendations for practice, and topics for future research along the lines of this campaign.

BACKGROUND

In this section we summarize succinctly a core set of literature from Educational Psychology and media richness. Although military operations appear to be set far apart from classroom education, the fundamental aspects of how people work, communicate and learn in virtual and physical environments spans such domains clearly. The connection between network-mediated interactions and media richness should be obvious.

Educational Psychology

The question of what can be accomplished via virtual versus physical environments has been addressed for many years by researchers in education, particularly where one or more information technologies (e.g., videotaping, TV broadcasting with telephone, video teleconferencing, web-based instruction) are used to enable distance education beyond the classroom. Historically there has been a strong bias toward classroom teaching (i.e., physical environment) and against distance education (i.e., virtual environment). Indeed, several governing bodies of higher education (e.g., accreditation boards) have issued blanket bans of distance education (Alberts, 1996, pp. 19-20), and where standards and policy organizations (e.g., Western Cooperative for Educational Telecommunication, Higher Education and Policy Council of the American Teachers Federation, Institute for Higher Education Policy) have sought to bolster distance education through quality standards, they have tended to focus on *minimum standards* to equate distance with classroom education (Alberts, Garstka, Hayes, & Signori, 2001).

However, numerous studies (Alberts, Garstka, & Stein, 1999) have compared the efficacy of classroom versus distance education, and most cases show no significant differences (Alberts & Hayes, 1995). Indeed, Bates and Poole (Alberts, 1996) indicate that "… the research evidence indicates clearly that technology-based teaching can be just as effective as face-to-face teaching" (p. 19). They go further by noting how technology enables some teaching techniques that are infeasible in the classroom, and they suggest that in some respects distance education can be *better* than its classroom counterpart (p. 23). This complements the group decision support systems literature (Alberts & Moffat, 2006, Ch. 7), which has shown for two decades that some aspects of group performance (e.g., mitigating rank and status differences, overcoming shyness and language difficulties, developing higher quality work products, examining a more complete range of alternatives and perspectives) are indeed *better* when people's interactions are mediated technologically than in face-to-face interactions. This leads us to propose similarly that some aspects of military work may be accomplished *more effectively* through virtual environments than through their physical counterparts.

Hypothesis 1a. The efficacy of military activities performed through virtual environments will exceed that of the same activities performed through physical environments.

Moreover, as technology advances and virtual environments become increasingly immersive (e.g., online video games and social networking sites, virtual and augmented reality, telepresence; see Allard, 1995; Bates & Poole, 2003; Bolger, 1990; Broll, Ohlenburg, Lindt, Herbst, & Braun, 2006), the range of feasible learning and experiential effects continues to expand. This leads us to propose that the degree of immersiveness enabled by a virtual environment contributes toward this result as well.

Hypothesis 1b. The efficacy of some military activities performed through virtual environments will increase in proportion with the degree of immersiveness.

Media Richness

Daft & Lengel (1986) define media richness "as the ability of information to change understanding within a time interval. Communication transactions that can overcome different frames of reference or clarify ambiguous issues to change understanding in a timely manner are considered rich. Communications that require a long time to enable understanding or that cannot overcome different perspectives are lower in richness. In a sense, richness pertains to the learning capacity of a communication." Oral media such as face-to-face and the telephone are considered to have higher levels of media richness than written media, such as interoffice mail. Additionally, synchronous media such as telephone and live chat sessions are considered to have higher levels of media richness than asynchronous media such as interoffice email or electronic mail (Markus, 1994).

Media richness theory rests on the assumption that organizations process information to reduce uncertainty and equivocality (Daft & Lengel, 1986). Uncertainty is defined by Galbraith (1977) as "the difference between the amount of information required to perform the task and the amount of information already possessed by the organization." Equivocality is the existence of multiple and conflicting interpretations about an organizational situation, or confusion, disagreement and lack of understanding about a particular problem-solving event (Daft & Macintosh, 1981; Trevino, 1987; Weick, 1979). Considering Media Richness theory as a prescriptive model (Daft & Lengel, 1984; Daft & Lengel, 1986), high and low levels of media richness provide distinct advantages in terms of reducing either equivocality or uncertainty.

Daft & Lengel (1984) conclude that written or text based media that are low in media richness are preferred for unequivocal messages, while a face-to-face environment that is high in media richness is preferred for messages containing equivocality. Rice and Shook (1990) suggest that media low in richness, such as business letters, convey less of a social presence and are less effective in terms of reducing equivocality through bargaining, negotiation, and conflict resolution. They also suggest that media low in richness is able to reduce uncertainty through the exchange of facts and information.

Within the context of military operations, the level of uncertainty, equivocality and the resultant information processing needs vary depending on the operational environment. The

comparative efficacy of using communication media within a virtual environment that is low in media richness versus face-to-face interaction that is high in media richness remains somewhat unclear and controversial, however.

Indeed, the prescriptions from Educational Psychology above (especially that performance in virtual environments will exceed that in physical counterparts) conflict in part with those from Media Richness Theory (especially that performance in media-rich environments will exceed that in media-poor counterparts). This is the case in particular where virtual environments (e.g., with non-immersive, textual interfaces) are expected to reflect low media richness. Moreover, this question and theoretical conflict take on increasing importance now, as many militaries are turning increasingly toward network-centric operations, through which most interactions between people are mediated by relatively non-immersive (esp. textual), media-poor environments. This leads to a conflicting hypothesis that can be tested empirically.

Hypothesis 2. The efficacy of some military activities performed under conditions of high media richness will exceed that of the same activities performed under conditions of low media richness.

RESEARCH DESIGN

In this section, we draw heavily from (Leweling & Nissen, 2007; Powley & Nissen, 2009; Powley & Nissen, 2009) to summarize the research design. Building upon prior experimentation, we employ the ELICIT multiplayer intelligence game in a laboratory setting to conduct a series of experiments and examine how virtual versus physical environment affects performance in the context of a counterterrorism problem solving task environment. We begin by describing this task environment and then outline the participants, groups, protocols, controls, manipulations and measurements used for experimentation in the virtual environment case. We close this section with a summary of the physical environment case.

Task Environment

The task requires a team of participants performing the roles of intelligence analysts to collaborate and identify a fictitious and stylized terrorist plot. The fictitious terrorist plot is described through a set of informational clues called “factoids” that have been developed systematically. The game’s design is similar to the Parker Brothers’ board game “Clue” in that it requires each player to analyze clues and combine assessments with other players to identify key aspects of the fictitious plot. Each factoid describes some aspect of the plot, but none is sufficient to answer all of the pertinent questions (i.e., Who will execute the attack? What is the target to be attacked? Where will the attack take place? When will the attack take place?).

The factoids are distributed among the players in a series of steps: each player receives two clues initially, followed by one after five minutes of play and another after ten minutes have elapsed. The factoid distribution is designed so that no single player can solve the problem individually and that the team of players cannot solve the problem until after the final distribution. In other words, the players must collaborate to solve the problem, and they are required to do so for a minimum of ten minutes. Evidence from previous experiments (e.g.,

Parity Communications Inc., 2006) suggests that play requires substantially more time (e.g., an hour or more).

Participants play the game in one or two modes. 1) For the virtual environment, they play via ELICIT client applications on separate computer workstations linked to a Web game server. Each subject has access to a set of five functions supported by the client application: 1) List, 2) Post, 3) Pull, 4) Share, and 5) Identify. After the game has completed, the administrator ends the simulation from the server application. The ELICIT application captures time-stamped interactions (e.g., Pose, Pull, Identify, List functions) including, for instance, when and which factoids are distributed to each player, when and which factoids are posted to which common screens, when and which common screens are viewed by each player, when and which factoids are shared between each player, and the time stamped results of each player's Identify attempt (i.e., to identify the who, what, where and when). 2) For the physical environment, they play through face-to-face interaction in rooms equipped with tables and white boards. Factoids are time stamped and distributed to the players on pieces of paper, and paper "postcards" are time stamped and used to collect players' Identify attempts. We do not attempt to log or time stamp players' other information sharing and processing activities.

The game requires considerable cognitive and collaborative effort to play well (i.e., identify the pertinent details of a terrorist plot), but experience indicates that such effort is within the capabilities of many people and groups.

Participants

Participants for this study are comprised of PhD students and faculty at a major US university. All participants have undergraduate college degrees and have completed graduate work at the masters and PhD levels. Most participants have direct military service as well, and some of the participants have worked professionally in military or government intelligence organizations. Hence the participants are representative in part of the kinds of relatively well-educated and experienced people who serve as professional intelligence analysts, particularly in national intelligence agencies.

In this experiment, participants are organized hierarchically as delineated in Figure 1. Such organization stratifies them into three functional levels. The Senior Leader is responsible for the intelligence organization as a whole and has four Team Leaders (middle managers) reporting directly. Each team leader in turn has three Team Members (Operators) reporting directly and is responsible for one set of details associated with the terrorist plot. For instance, Team Leader (Who) and his or her team are responsible for the "who" details (e.g., which terrorist organization is involved) of the plot, Team Leader (What) and his or her team are responsible for the "what" details (e.g., what the likely target is), and so forth for "where" and "when." Participants are assigned randomly to these roles.

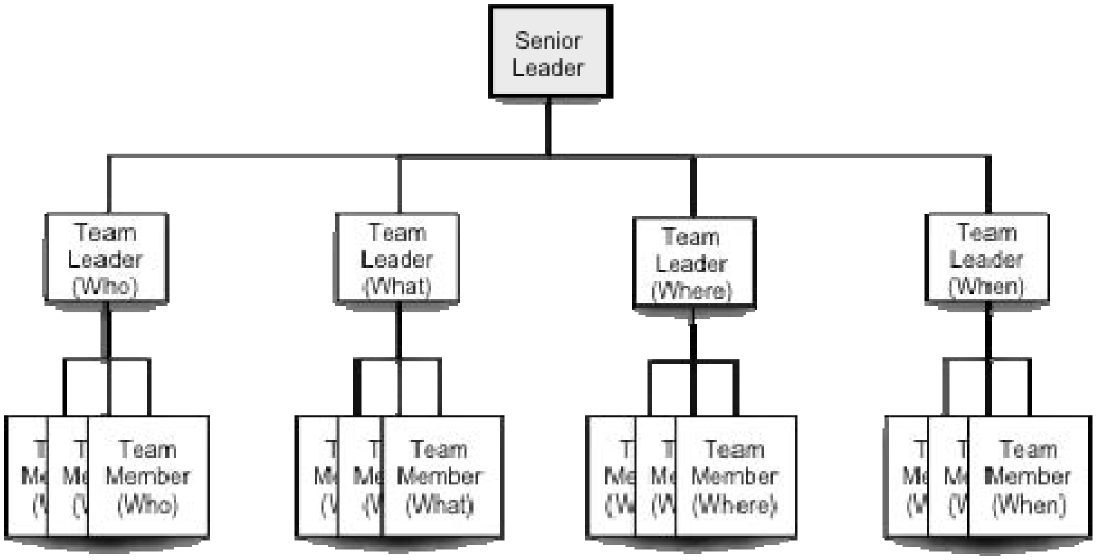

Figure 1 Hierarchical Organization

Treatment Groups

Participants are assigned to play in each of two groups of the same 17 members (i.e., one to play in the virtual environment, one to play in the physical environment). To address concerns with learning and bias, participants are assigned randomly to the 17 different roles in each game, and the order of the two groups (i.e., virtual or physical) is decided randomly; in this experiment, participants play in the virtual environment first and then the physical environment. Additionally, different yet structurally equivalent versions of the game are played each time, and the two experiment sessions are conducted two weeks apart; this reduces learning opportunities yet ensures that both problem-solving and information-sharing tasks are comparable. We summarize protocols and manipulations for each group below.

Virtual Environment. Protocols and manipulations for the virtual environment are designed to be very consistent with most prior experiments using ELICIT. Participants report to a networked classroom on their assigned day for the experiment. Once seated, participants are allotted ten minutes to read a set of instructions pertaining to both the experiment and the ELICIT environment; they are encouraged to ask questions about the experimental settings and environment. Once participants read the instructions they have ten minutes to discuss their approach to the problem-solving scenario with others in their group and take a short break before beginning. The first five minutes are allotted to discussion between the Senior Leader and four Team Leaders; the discussion is via e-mail, and Operators do not participate. The next five minutes are allotted to separate discussions between each Team Leader and his or her three Operators; the discussions are via e-mail also, and each team conducts them separately.

Once the game begins, each participant receives unique factoids in three phases: 1) two factoids initially when the game begins, 2) one after five minutes, and 3) one at the ten-minute mark. Role-specific factoids are distributed automatically by ELICIT and in a manner ensuring:

a) that no player can solve the plot alone, and b) that the plot cannot be solved until all factoids have been distributed. Factoids are time stamped and appear automatically on players' ELICIT screens, after which they can be posted, pulled and shared. As noted above, all Post, Pull, Share and Identify actions are time stamped and logged by the game server.

Participants communicate with one another during game play using only the computer-network capabilities supported by ELICIT (esp. Post, Pull and Share) and readily available network communication technologies (esp. e-mail); no verbal communication is allowed. This represents a slight departure from most prior experiments, in which e-mail and like technologies (DoD, 1980) were not permitted. Prior experiments appear to have suffered from reduced external validity due to uncharacteristic constraints placed on communications, and the ubiquity of e-mail suggests that this modification will increase such validity.

Another slight departure from prior experiments pertains to the pseudonyms used in the game: each player is assigned randomly a pseudonym associated with his or her role in the game, and to conceal players' identities in previous experiments, their pseudonyms were not revealed. This proved to be an unenforceable protocol in many prior experiments, as players were situated often in the same room and able to ascertain one another's pseudonyms. Alternatively, in this experiment players are informed explicitly of one another's pseudonyms. This helps to overcome the limitation of prior experiments, and it facilitates experiments in the physical environment where players' identities would be very difficult to conceal.

Additionally, the ELICIT software limits participants' Post (i.e., sharing factoids with others) and Pull (accessing factoids posted by others) access to specific common screens within this manipulation. Specifically, those players in the "who" group, for instance, are allowed to Post to and Pull from only one of the four common screens (i.e., the "who" screen) noted above. The only exception applies to the Senior Leader, who has post-pull access to all four common screens. This is comparable to most prior experiments.

Alternatively, a third slight deviation from prior experiments involves the ability of players to share factoids with one another. Most prior experiments enable every player to share with any other player, regardless of team assignment. This appears to conflict with the post and pull restrictions from above and to mitigate the effects of hierarchical organization. In this experiment Operators are permitted to share only with members of their own teams, and Team Leaders are permitted to share only with Operators on their respective teams in addition to the Senior Leader. Further, the Senior Leader communicates only with Team Leaders. This reinforces hierarchical communication and chain of command.

The simulation ends after approximately 45 minutes. All players are given the option to Identify the plot details if they have not done so already. Players are instructed to Identify only once during game play, and they are incentivized both to solve the plot individually as well as to collaborate so that others on the team (esp. the Senior Leader) solve the plot quickly and accurately.

As a note, in this experiment we are able to examine only a single level of immersiveness: that enabled via the conventional ELICIT interface. Although we are working in parallel to enable more immersive environments, such interfaces are not available at the time of this experiment; hence we must defer our testing of Hypothesis 1b to the next experiment.

Physical Environment. Protocols and manipulations for the physical environment are designed to be as consistent as possible with those summarized above for the virtual environment. Participants report to a classroom with desks, tables and whiteboards on their assigned day for

the experiment. Once seated, participants are allotted ten minutes to read a set of instructions pertaining to both the experiment and the physical environment; they are encouraged to ask questions about the experimental settings and environment. Once participants read the instructions they have ten minutes to discuss their approach to the problem-solving scenario with others in their group and take a short break before beginning. The first five minutes are allotted to discussion between the Senior Leader and four Team Leaders; the discussion is via face-to-face conversation, and Operators do not participate (i.e., they leave the room). The next five minutes are allotted to separate discussions between each Team Leader and his or her three Operators; the discussions are via face-to-face conversation also, and each team conducts them separately (i.e., they meet in separate rooms).

Once the game begins, each participant receives the same unique factoids in the same three phases summarized above: 1) two factoids initially when the game begins, 2) one after five minutes, and 3) one at the ten-minute mark. In this case, however, role-specific factoids are distributed manually on pieces of paper. As above, the factoids are distributed in a manner ensuring: a) that no player can solve the plot alone, and b) that the plot cannot be solved until all factoids have been distributed. The time of distribution is noted on each factoid. Unlike in the virtual environment above, we make no attempt in this physical environment to record or note the time when factoids are posted (e.g., on the white board) pulled (e.g., by a player taking notes from the white board) or shared (e.g., verbally). However, consistent with the virtual environment above, the time of each Identify action is recorded.

In great contrast with the virtual environment, participants communicate with one another during game play using only face-to-face conversation and the white board; no computer-mediated communication is allowed. The only exception involves the telephone; after the teams move to geographically separate rooms, Team Leaders are allowed to speak by telephone with the Senior Leader. Telephony is a century-old technology that we do not consider part of a virtual environment.

As above, each player is assigned randomly a pseudonym associated with his or her role in the game, and in this experiment players are informed explicitly of one another's pseudonyms. This helps to overcome the limitation of prior experiments, and it facilitates experiments in the physical environment where players' identities would be very difficult to conceal. Likewise, players' physical separation in different rooms limits participants' Post (i.e., sharing factoids with others) and Pull (accessing factoids posted by others) access to specific white boards in each of the rooms within this manipulation. Specifically, those players in the "who" group, for instance, are allowed to Post to and Pull from only one of the four common white boards (i.e., in the "who" room) noted above. The only exception applies to the Senior Leader, who has access to all four rooms. This preserves the limitations imposed in the virtual environment above.

Also as above, in this experiment Operators are permitted to share only with members of their own teams, and Team Leaders are permitted to share only with Operators on their respective teams in addition to the Senior Leader. Further, the Senior Leader communicates only with Team Leaders. This reinforces hierarchical communication and chain of command.

The simulation ends after approximately 45 minutes. All players are given the option to identify the plot details if they have not done so already. Players are instructed to Identify only once during game play, and they are incentivized both to solve the plot individually as well as to collaborate so that others on the team (esp. the Senior Leader) solve the plot quickly and accurately.

Measurements

Following (Leweling & Nissen, 2007), we operationalize performance as a two-dimensional dependent variable comprised of: 1) *speed* (i.e., time to identify plot details correctly) and 2) *accuracy* (i.e., correct identification of plot details). These dependent measures are informed by literature in the psychological and organizational domains that suggest a trade-off exists between time and accuracy in tasks requiring high cognition and/or advanced motor skills (e.g., see Beersma et al., 2003; Elliott, Helsen, & Chua, 2001a; Guzzo & Dickson, 1996; Meyer, Irwin, Osman, & Kounios, 1998; Plamondon & Alimi, 1997; Rogers & Monsell, 1995) at both the individual and team/group levels of analysis.

In the first component, speed pertains to how long it takes a participant to submit his or her identification of the terrorist plot details. For ease of comparison, the scale for this speed measurement is normalized to a 0-1 scale, with 1 being more desirable (i.e., faster). Measuring and normalizing time is straightforward, as the time for each participant's identification is logged to the nearest second by the software. Specifically, each participant's elapsed time is recorded when he or she uses ELICIT to Identify the plot. To construct a scale in which faster speeds (i.e., shorter times to Identify) result in larger values, a baseline time is established as the maximum time required for the slowest of all participants (i.e., 3000 seconds in this experiment). Each participant's time to identify is related to this baseline and normalized to produce a scaled score according to the formula: speed = (3000 – time) / 3000; that is, an individual participant's time (say, for example, 2375 seconds) would be converted to a speed score as: speed = (3000 – 2375) / 3000 = 0.2083. All participants' times are converted to speed scores in this same manner and using this same baseline.

The second component of performance, accuracy, refers to the quality of the identification of the impending terrorist attack (i.e., Who, What, Where, and When). Each participant's Identify action is scored with a value of 1 for each correct answer to the Who, What and Where aspect of the solution. Note, however, that the When aspect of the solution includes three components (i.e., Month, Day, and Time). In order to avoid weighting this aspect more heavily than the other three, each participant's Identify action is scored with a value of 1/3 for each correct answer. The resulting sum is divided by four to construct a [0-1] scale; that is, an individual participant's Identify (say, for example, identifies the Who, What and Where aspects correctly but is correct only on the day and not the month or time components of the When aspect) would be converted to an accuracy score as: accuracy = (1 + 1 + 1 + 1/3) / 4 = 0.83.

RESULTS

In this section, we summarize results from the experiment, beginning with an overview of key statistical results and followed by discussion of their key implications.

Statistical Results

Table 1 presents the descriptive statistics for the key variables in our analysis. To communicate our results in clearly understandable units, we present the speed scores in seconds, as opposed to scaled scores. The speed scores can easily be recalculated as described above. In the table, "CM" refers to computer-mediated, virtual environment, and "FTF" refers to face-to-face, physical environment. We analyze the accuracy of identification as a composite score as discussed above,

but we also present the results of analyzing separately each component of the accuracy identification. To enhance interpretation, we use **bold font** to highlight the higher performing score in each CM-FTF contrast.

TABLE 1

Descriptive Statistics

Variable	Mean	N	Standard Deviation	Standard Error
ID Time CM (in Seconds)	2685	14	219	58
ID Time FTF (in Seconds)	**2554**	14	279	74
Who Score CM	.140	14	.363	.097
Who Score FTF	**.790**	14	.426	.114
What Score CM	.321	14	.249	.066
What Score FTF	**.536**	14	.365	.098
Where Score CM	.570	14	.514	.137
Where Score FTF	**.790**	14	.426	.114
When Score CM	.262	14	.297	.079
When Score FTF	**.333**	14	.320	.086
Identify Composite CM	.324	14	.206	.055
Identify Composite FTF	**.610**	14	.327	.087

To summarize, the mean identification speed appears faster in the physical environment (2554 versus 2685 seconds, roughly two minutes faster), but the standard deviation is greater (279 versus 220 seconds, roughly one minute). Likewise, the mean accuracy scores appear larger (i.e., more accurate) in the physical environment for every component of the solution (i.e., who, what, where, and when) as well as for the overall composite identification score, and as above, the standard deviations are greater in nearly every component as well. Participants in the physical environment appear to outperform their counterparts in the virtual environment, but variability of such performance is greater.

Table 2 presents the results of the hypotheses tests. Our first hypothesis posits that the efficacy of military activities performed through virtual environments will exceed that of the same activities performed through physical environments. Our second hypothesis, essentially the

competing position from our first hypothesis, posits that the efficacy of some military activities performed under conditions of high media richness (our physical environment) will exceed that of the same activities performed under conditions of low media richness (our virtual environment).

Because the same players participated in both experiment sessions, we employ the two-tailed t-test of paired samples to take advantage of the blocking and increase statistical power. As in the table above, “CM” refers to computer-mediated, virtual environment, and “FTF” refers to face-to-face, physical environment. We also continue our practice of presenting the speed results in actual time (seconds) for ease of interpretation, while noting that converting to the standardized scores we mentioned is straightforward using the formula above.

Each row of the table summarizes a contrast between CM and FTF. For instance, the first row summarizes the difference between identification time (in seconds) for the CM and FTF sessions (i.e., CM time minus FTF time); we report the mean difference (131 seconds), t value (1.242), degrees of freedom (13) and significance (.236, 2-tailed test).

TABLE 2

Results of Hypotheses Tests

	Statistical Support			
Variable	**Mean Value**	**t**	**df**	**Significance (2-tailed)**
ID Time CM - ID Time FTF in Seconds	131	1.242	13	.236
Who Score CM - Who Score FTF	-.643	-4.837	13	**.000**
What Score CM - What Score FTF	-.2143	-1.578	13	.139
Where Score CM - Where Score FTF	-.214	-1.000	13	.336
When Score CM - When Score FTF	-.071429	-.612	13	.551
CM Composite – FTF Composite	-.286	-2.362	13	**.034**

Notice that the difference in identification speeds is not statistically significant ($p = 0.236$), so we are unable to support either of our competing hypotheses in terms of the speed measure. Likewise, most of the accuracy contrasts are not significant either; the “who” component is highly significant ($p < .001$), and the composite accuracy score reflects considerable significance ($p < .05$) too. This provides some support for our second hypothesis: the media-rich environment supports more accurate performance.

Key Findings

The results summarized above contain two important findings for organizational leaders seeking to better understand the comparative performance of people working in media-rich, face-to-face environments versus textual, (low-immersive) virtual counterparts. The results of this study support media richness theory, while at the same time they provide a plausible explanation for the theoretical conflict between the Educational Psychology literature and media richness theory.

As summarized above, performance in the virtual environment is worse than in the physical environment. This provides support *against* hypothesis one: *The efficacy of military activities performed through virtual environments will exceed that of the same activities performed through physical environments.*

Virtual environments that are low in media richness, such as the one used in this experiment and those commonly found in business and military organizations, provide little opportunity for intelligence professionals to reduce task equivocality, often resulting in reduced group accuracy. Although virtual environments can provide access to experts and accelerate information sharing among distributed decision makers, thus reducing uncertainty, such access does not appear to be compelling in this experiment.

Alternatively, at least in terms of accuracy, performance in the media rich, face-to-face environment is better than in the low-immersive virtual environment. This provides support *for* hypothesis two: *The efficacy of some military activities performed under conditions of high media richness will exceed that of the same activities performed under conditions of low media richness.*

In this experiment, speed appears to be relatively insensitive to physical or virtual environment, but such environmental choice affects accuracy. This calls in question what appears to be a strong assumption reflected in counterterrorism intelligence organizations: that physically distributed, virtual environments should be employed. Such environments do not appear to gain a speed advantage, yet they do appear to suffer an accuracy disadvantage. This suggests that intelligence leaders and policy makers may benefit by rethinking their organizational assumptions, particularly where accuracy is important.

Moreover, only one dimension of *accuracy* (i.e., “Who”) reflects statistical significance in this experiment, yet this dimension appears to have sufficient influence to make the composite accuracy difference statistically significant as well. Hence, of the four accuracy components, intelligence organizations would appear wise to focus their attention and resources on identifying *who* is planning terrorist attacks, for this appears to have strong influence on overall counterterrorism intelligence efficacy.

Further, considering this experiment, some of the theoretical conflict between the Educational Psychology literature and the media richness literature may be resolved by exploring the variation in task uncertainty and task equivocality among tasks. Educational tasks focused on a particular set of learning objectives tend to be static, well defined, and operate within a pedagogical framework. Military and business tasks that are focused on sense and response activities within an emergent environment are dynamic, often based on interpretation, and more closely align with the principles of andragogy.

Within an educational setting where task equivocality is low, it therefore makes sense how virtual environments that are low in media richness can out-perform face-to-face environments by providing efficient real-time access to expert knowledge and information that reduces uncertainty in the task environment. We gain further insight into how military and

business organizations, operating within emergent, sense-and-respond ecospaces where task equivocality is high, can benefit more from face-to-face environments than virtual environments through greater shared and more accurate interpretation of the task environment. Hence a plausible resolution for the theoretical conflict between the Educational Psychology literature and media richness theory resides in defining the task in terms of the level of task uncertainty and task equivocality.

CONCLUSION

Research in command and control is advancing rapidly through a campaign of laboratory experimentation using the ELICIT (Experimental Laboratory for Investigating Collaboration, Information-sharing, and Trust) multiplayer online counterterrorism intelligence game. In most ELICIT experiments, participants play the game through a Web interface and interact with one another solely through textual information exchange. This mirrors in large part the network-centric environment associated with most counterterrorism intelligence work in practice, and it reflects what appears to be a widespread assumption about how to organize such work: in a physically distributed, virtual manner.

Such reflection is consistent with considerable research (e.g., in Educational Psychology) prescribing such distributed, virtual environments for work performance. Alternatively, substantial research (e.g., Media Richness Theory) suggests instead that a more personal, physical environment offers potential to improve performance. Hence we have a theoretical conflict with potential to affect how the important work of counterterrorism intelligence is organized.

The research described in this article addresses this theoretical conflict through a series of experiments to assess the comparative performance of people working in physical, face-to-face versus textual, virtual environments. In particular, we employ the ELICIT multiplayer counterterrorism intelligence game to examine the comparative performance of people accomplishing counterterrorism intelligence tasks across such different environments. Exercising great care to match experiment conditions and control for factors other than task environment, we set up and run a series of experiments to develop empirical evidence. The same group of participants play the ELICIT game via a virtual environments and then via its physical counterpart.

Results elucidate important comparative performance effects. Specifically, we find no statistical significance in performance differences measured in terms of speed (i.e., how quickly a team can identify a terrorist plot), but differences in terms of accuracy (i.e., how many plot details can be identified correctly) are statistically significant overall: participants in the physical environment outperform counterparts in the virtual environment. Hence—where accuracy is concerned—intelligence leaders and policy makers may find cause to rethink the widespread organizational assumption that favors physically distributed, virtual work.

Further, through consideration of how uncertainty and equivocality affect the kinds of environments associated generally with educational contexts versus military and business processes, results of this experiment offer some resolution of the theoretical conflict between Educational Psychology and Media Richness Theory: where uncertainty prevails (e.g., in the educational context), virtual environments may be adequate and even outperform their physical counterparts, but where equivocality predominates (e.g., military and business processes), richer

media offered through physical environments appears to be important and to support superior performance.

These results suggest compelling follow-on experiments as well as practical implications. For instance, as noted above, we are unable to examine *immersiveness* as what is hypothesized to be an important variable in terms of virtual environments. Indeed, this experiment includes only the single, relatively media-poor, textual environment supported by the ELICIT interface. Work to develop more immersive environments (e.g., where participants play the game via avatars) may provide opportunities to extend this series of experiments to examine explicitly the impact that immersiveness has on performance.

As another instance, we were unable to examine the *Edge Organization* as an alternative to the Hierarchy. Indeed, this experiment includes only the single, common, hierarchical organization structure in both the physical and virtual environments. Work to conduct additional experiments that examine comparative performance—in both physical and virtual environments—of hierarchical versus edge organization may enrich this investigation greatly. Moreover, these two experiment extensions can be combined for even greater impact: experiments to examine the interactions between immersiveness and organization offer potential to yield highly informative results. We are working at present on such extensions and results.

In terms of practical implications, these results call into question the predominate, physically distributed, textual/virtual environment employed for counterterrorism intelligence work. Intelligence leaders and policy makers may have cause to rethink their organizing assumptions, particularly where accuracy in terms of identifying terrorist plots is important. Although much intelligence information is collected near the sources and locations of terrorist plotters and suspects—and hence must be physically distributed by necessity—the *analysis* of such intelligence information does not have the same necessary cause for physically distributed work, and counterterrorism intelligence leaders and policy makers may find it useful to collocate analysts in physical environments that afford media-rich, face-to-face interactions.

As with any study, the series of experiments reported in this article have limitations. For instance, we note above how we are unable to examine different degrees of immersiveness or contrast alternate organizational forms. The literature suggests that such variables are important; hence our results should be interpreted as partial until the complementary experiments noted above have been completed.

As another instance, the experiment itself had a couple of unplanned events occur that may affect the results. For one, although we permitted telephonic communication in the physical-environment experiment sessions, *no one used the telephone* during such sessions. It is unclear how telephone use may have affected the results—and in many respects this reflects a purer physical-virtual environment contrast—but as with most experiments, not everything went exactly as planned.

For another, technical issues precluded participation of three ELICIT roles; only 14 of the 17 planned roles were played in the virtual environment session. This clearly limited the extent to which information could be shared and processed, and overall performance likely suffered as a result. For consistency with the physical environment session, only these same 14 roles were played. Hence our comparative results reflect the exact same situation in terms of role playing and opportunities for information sharing and processing, and overall performance likely suffered as a result in both experiment sessions. However, the extent to which performance in physical versus virtual environments is sensitive to the absence of participants in key roles (e.g.,

robust to missing intelligence analysts and information) remains unclear. This reflects another opportunity for future research.

Many other opportunities for research along the lines of this investigation obtain as well. For one, a parallel article (Elliott, Helsen, & Chua, 2001b) reports results of developing a set of semi-intelligent software agents to play the ELICIT game in a manner reflecting human performance; that is, using ELICIT (Cacioppe, 1999) to enable software agents to play the game, such agents are tailored specifically to reflect the information sharing and processing behavior of the 14 participants in our experiment, and the comparative performance of human versus software agents is examined. This offers even greater opportunity to integration with the kinds of future experiments noted above.

For instance, software agents can play—either in conjunction with or in lieu of human participants—across a range of virtual environments and organizational forms. Specifically, we may be able to adapt the software agents to play in highly immersive virtual environments as well as via the standard, textual interface supported by ELICIT. We may likewise be able to adapt them to play in edge forms as well as hierarchies, and the whole aspect of software agents performing in support of versus instead of humans elucidates an exciting, highly underexplored avenue for future research.

Although the research described in this article provides only one metaphorical step in this direction, it is an important step, one that challenges prevailing organizing wisdom in terms of counterterrorism intelligence work and that offers to reconcile theoretical conflict pertaining to performance in physical versus virtual environments. We look forward to continuing along the lines of this investigation, and we welcome others to join and complement our work.

REFERENCES

Alberts, D. S. (1996). *The unintended consequences of information age technologies*. Washington, DC: CCRP.

Alberts, D. S., Garstka, Hayes, & Signori. (2001). *Understanding information age warfare*. Washington, DC: CCRP.

Alberts, D. S., Garstka, & Stein. (1999). *Network centric warfare*. Washington, DC: CCRP.

Alberts, D. S., & Hayes, R. E. (1995). *Command arrangements for peace operations*. Washington DC: CCRP.

Alberts, D. S., & Moffat, J. (2006). *Maturity levels for NATO NEC command*. UK: UK Defence Science & Technology Laboratory.

Allard, K. C. (1995). In Institute for National Strategic Studies (Ed.), *Somalia operations: Lessons learned* Institute for National Strategic Studies.

Bates, A. W., & Poole, G. (2003). *Effective teaching with technology in higher education: Foundations for success*. San Francisco, CA: Jossey-Bass.

Beersma, B., Hollenbeck, J. R., Humphrey, S. E., Moon, H., Conlon, D. E., & Ilgen, D. R. (2003). Cooperation, competition, and team performance: Toward a contingency approach. *Academy of Management Journal, 46*(5), 572-590.

Bolger, M. D. P. (1990). Command or control. *Military Review*, (July), 69-79.

Broll, W., Ohlenburg, J., Lindt, I., Herbst, I., & Braun, A. (2006). Meeting technology challenges of pervasive augmented reality games.

Cacioppe, R. (1999). Using team – individual reward and recognition strategies to drive organizational success. *Leadership and Organization Development Journal*, *20*, 322-331.

Daft, R. L., & Lengel, R. H. (1984). Information richness: A new approach to managerial behavior and organizational design. *Research in Organizational Behavior, 6*, 191-233.

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science, 32*(5, Organization Design), 554-571.

Daft, R. L., & Macintosh, N. B. (1981). A tentative exploration into the amount and equivocality of information processing in organizational work units. *Administrative Science Quarterly, 26*(2), 207-224.

DoD. (1980). *Proceedings for quantitative assessment of utility of command and control systems* No. MTR-80W00025)

Elliott, D., Helsen, W. F., & Chua, R. (2001a). A century later: Woodworth's (1899) two-component model of goal directed aiming. *Psychological Bulletin, 127*(3), 342-357.

Elliott, D., Helsen, W. F., & Chua, R. (2001b). A century later: Woodworth's (1899) two-component model of goal directed aiming. *Psychological Bulletin, 127*(3), 342-357.

Galbraith, J. R. (1977). *Organization design*. Reading, Mass: Addison-Wesley Pub. Co.

Guzzo, R. A., & Dickson, M. W. (1996). Teams in organizations: Recent research on performance and effectiveness. *Annual Review of Psychology*, *47*(1), 307.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York, NY: Cambridge University Press.

Leweling, T. A., & Nissen, M. E. (2007). Hypothesis testing of edge organizations: Laboratory experimentation using the ELICIT multiplayer intelligence game. *Proceedings International Command & Control Research & Technology Symposium,* Newport, RI.

Markus, M. L. (1994). Electronic mail as the medium of managerial choice. *Organization Science, 5*(4), 502-527.

Meyer, D. E., Irwin, D. E., Osman, A. M., & Kounios, J. (1998). The dynamics of cognition and action: Mental processes inferred from time-accuracy decomposition. *Psychological Review, 95*, 183-237.

Parity Communications Inc. (2006). *Experiments in command and control within edge organizations: Final report* No. EBR06-0002). Boston, MA:

Plamondon, R., & Alimi, A. M. (1997). Time/Accuracy trade-offs in target-directed movements. *Behavioral and Brain Sciences, 20,* 279-303.

Powley, E. H., & Nissen, M. E. (2009). Trust-mistrust as a design contingency: Laboratory experimentation in a counterterrorism context. Washington, DC.

Ramo, J. C. (2009). *The age of the unthinkable: Why the new world disorder constantly surprises us and what we can do about it*. New York: Little, Brown and Company.

Rice, R. E., & Shook, D. E. (1990). Relationships of jop categories and organizational levels to use of communication channels, including electronic mail: A meta-analysis and extension. *The Journal of Management Studies, 27*(2), 195.

Rivkin, J. W. (2000). Imitation of complex strategies. *Management Science, 46,* 824-844.

Roger B. Mason, University of Surrey,Surrey, UK. (2007). The external environment's effect on management and strategy. *Management Decision, 45*(1), 10.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology, 124,* 207-231.

Trevino, L. K. (1987). Message equivocality, media selection, and manager performance: Implications for information systems. *MIS Quarterly, 17,* 355.

Weick, K. E. (1979). *The social psychology of organizing* (2nd ed.) Addison-Wesley.